LETTER TO THE EDITOR

A clinical study of preimplantation DNA methylation screening in assisted reproductive technology

Cell Research (2023) 0:1–3; https://doi.org/10.1038/s41422-023-00809-z

Dear Editor,

Assisted reproductive technology (ART) has been used for more than 45 years, by which more than 9 million infants were born. One of the key factors to determine the success rate of the ART is embryo selection.[1,2] Due to the existence of aneuploidy in 20%–30% of blastocysts, preimplantation genetic screening (PGS) has been widely used to increase live birth rate. However, the live birth rate remains around 50% even with the help of PGS.[2] DNA methylation is known to play an important role during embryogenesis.[3–5] A previous study showed that a large proportion of human embryos have abnormal DNA methylome, and indicated that preimplantation DNA methylation screening (PIMS) can analyze both copy number variation (CNV) and global DNA methylation level.[5] However, whether DNA methylation patterns can affect the clinical outcome of ART has not been investigated in clinics. In this regard, we performed a clinical trial of PIMS (trial number: NCT03642574). We aimed to examine the relationship between embryo methylome and the clinical outcome of ART.

182 families including 800 blastocysts were enrolled in PIMS. 3–5 biopsied cells from trophectoderm of each blastocyst were measured with whole genome bisulfite sequencing. Using methylome data, we analyzed CNV and global average methylation levels (Supplementary information).[5] The global methylation level is the average of all sequencing-covered CpGs. Our data show that the methylation level variation of different trophectoderm cells in the same embryo is similar (Supplementary information, Fig. S1a). Since PIMS can simultaneously provide information on CNV and methylation, there is no need to perform PGS to analyze CNV anymore. Not knowing what kind of methylome can produce the best clinical outcome of ART, we selected embryos only based on CNV instead of DNA methylome information. In total, 163 euploid embryos underwent elective single embryo transfers (Fig. 1a; Supplementary information, Fig. S2a), and 3 cases of twin pregnancy were excluded from the downstream analysis. The clinical data show 57 pregnancy failures, 13 pregnancy losses and 90 (56.25%) live birth neonates (Supplementary information, Table S1, S2).

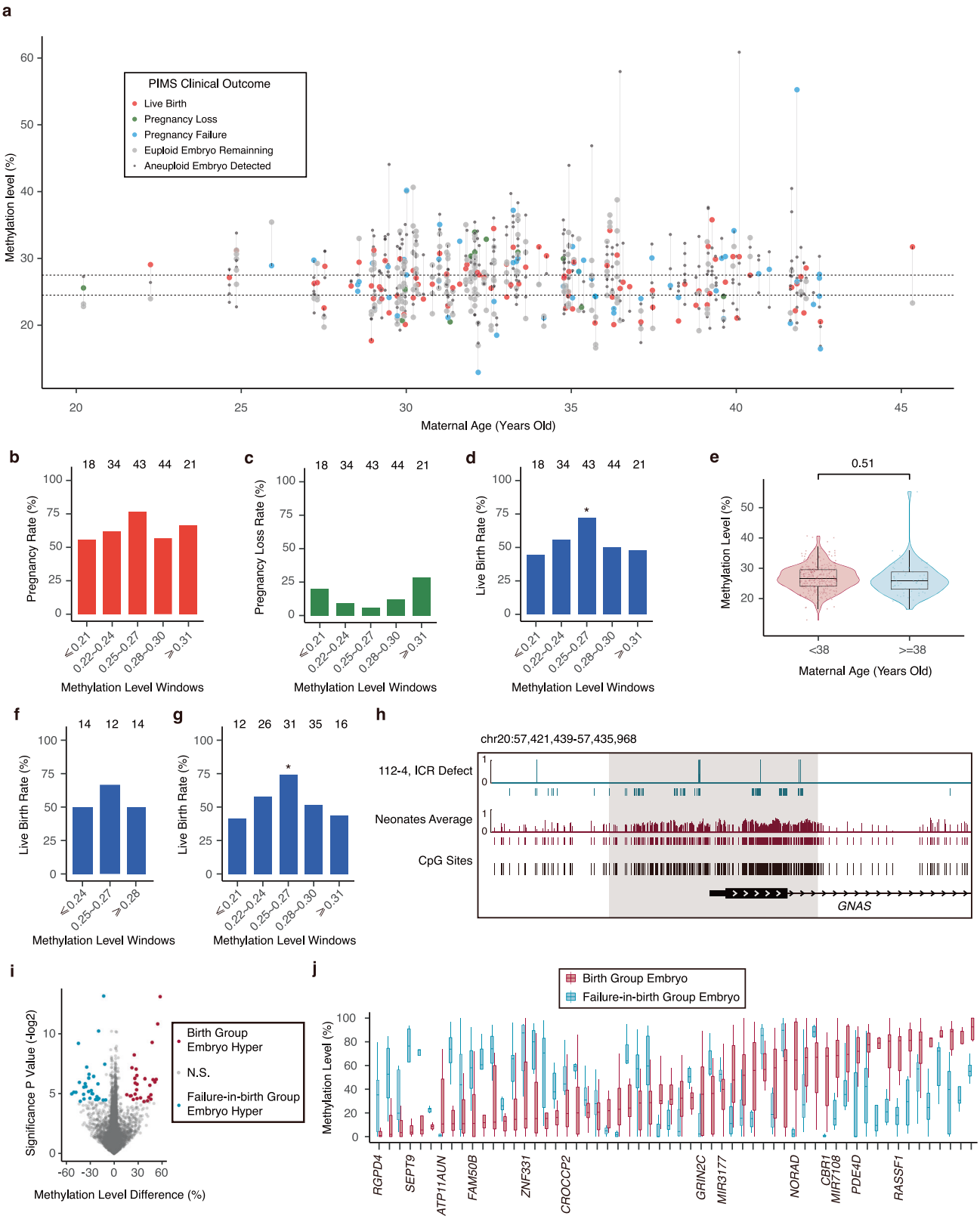
To study whether embryos with different methylation levels have different clinical outcomes, we assigned embryos into different groups according to their DNA methylation level. Notably, the embryos with DNA methylation levels between 0.25 and 0.27 produce significantly higher live birth rates than the embryos with other methylation levels (odds ratio [OR], 2.52; 95% confidence interval [CI], 1.13–5.95; $P = 0.02$). The pregnancy rate shows similar trends (OR, 2.21; CI, 0.95–5.51; $P = 0.06$), while the pregnancy loss rate shows the opposite trends (OR, 0.35; CI, 0.04–1.75; $P = 0.22$) (Fig. 1b–d). Overall, the higher difference of the methylation value from the window of 0.25–0.27 is, the lower the birth rate and pregnancy rate become (Fig. 1b, c); meanwhile, the pregnancy loss rate becomes higher also (Fig. 1d). Therefore, euploid embryos with DNA methylation level closest to 0.25–0.27 should be preferentially selected for transferring.

The proportion of aneuploid blastocysts in younger women is significantly less than that in women with advanced maternal age (AMA, ≥ 38 year old).[6] During ART practice, PGS can increase the live birth rate for AMA women, while it has a limited effect on younger women.[2,6,7] In contrast, the DNA methylation level variance of embryos is similar between both younger and AMA women (Fig. 1e). Further analysis shows that the association between DNA methylation level and the clinical outcome can be observed in both younger and AMA women (Fig. 1f, g; Supplementary information, Fig. S3a–d). We further checked the distribution of the age in each methylation level window, which shows there are no significant differences in maternal age among different windows (Supplementary information, Fig. S4a). These data indicate that the DNA methylation level affects the clinical outcome in both younger and AMA women.

DNA methylation abnormalities in imprinted control regions (ICRs) can cause imprinting disorders. The birth defect of imprinting disorders occurs in about 0.2% of the human population, and this rate is doubled in ART-born babies.[8–10] Unfortunately, imprinting disorders cannot be avoided during current ART practice. Here, we checked the methylation status of known germline ICRs.[11–13] As expected, all these germline ICRs are middle methylated. In addition, about half of the reads in ICRs are fully methylated reads, and about half reads are unmethylated (Supplementary information, Fig. S5a). Furthermore, our data show that some embryos have abnormal methylation states in germline ICR. For example, the methylation level of *GNAS* ICR in an unused embryo is unmethylated with a significant absence of hypermethylated reads (methylation level: 0.07, $P = 0.016$) (Fig. 1h). Our data suggest that we can use PIMS to exclude the embryos with methylation mutations in ICRs, which can potentially decrease the rate of imprinting disorders during ART practice.

We noticed that some embryos with methylation in the 0.25–0.27 window could not lead to live birth, suggesting that some important regions with abnormal methylation states might lead to the failure of live birth. To test whether the embryos can lead to live birth or not, we divided the embryos into live birth group and failure birth group. Fifty eight differential methylated regions (DMRs) between these two groups were identified with a *P* value of less than 0.05 (Fig. 1i). Our data show that 64 genes locate within 10 kb of these DMRs, and the promoters of 13 genes overlap with DMRs (Fig. 1j; Supplementary information, Table S3). Some of these genes are known to regulate embryo development and be affected by

Received: 15 January 2023 Accepted: 3 April 2023
Published online: 08 May 2023



DNA methylation. For example, DNA methylation of *SEPT9* promoter is associated with cervical cancer.[14]

Taken together, the PIMS method can examine both CNV and methylation information of embryos, so it can replace the method of preimplantation genetic testing for aneuploidy. Embryos with better methylation states can produce better clinical outcomes during ART. Therefore, PIMS can potentially increase the live birth rate of ART, and decrease the birth defect rate.

Fig. 1 The clinical outcomes of PIMS. a Overview of enrolled participants and the clinical outcomes of their embryos in PIMS. Embryos from the same participant were presented in the same column. Dashed lines indicate the methylation level window between 0.25–0.27. **b**–**d** Barplots showing pregnancy rate (**b**), pregnancy loss rate (**c**) and live birth rate (**d**) of the embryos with different methylation levels. **e** Methylation levels of the embryos for younger and AMA women. Significance of two-sided *t*-test was indicated above. **f**, **g** Barplots showing live birth rate of embryos with different methylation levels for younger (**f**) or AMA women (**g**). * refers to $P < 0.05$ and the total number of embryos transferred in each methylation level window was indicated in each column in (**b**–**d**, **f**, **g**). **h** Visual track of the embryos with unmethylated *GNAS* ICR ($P = 0.016$). Genomic (black) and sequencing covered (red or cyan) CpG sites were indicated by the vertical bars. The gray shaded box indicates *GNAS* ICR. **i** Differential methylated regions between the birth and failure groups. **j** Methylation level distribution of embryos in the birth and failure groups for the identified DMR in (**i**). Genes whose promoter overlapped with DMR were indicated in their corresponding columns in (**j**). Boxes and whiskers in (**e**) and (**j**) represent the 25th/75th percentiles and 1.5× interquartile range, respectively.

Yuan Gao[1,2,3,8] ✉, Lizhi Yi[4,5,6,8], Jianhong Zhan[4,5,8], Lijuan Wang[1,2,3,8], Xuelong Yao[6,8], Junhao Yan[1,2,3], Sijing Jian[4,5], Lei Gao[4,5], Mamadboqirova Farangez[4,5], Ming Gao[1,2,3], Yang Zou[1,2,3], Xuan Gao[1,2,3], Keliang Wu[1,2,3], Jiang Liu[4,5] ✉ and Zi-Jiang Chen[1,2,3,7] ✉

[1]Center for Reproductive Medicine, Shandong University, Jinan, Shandong, China. [2]Shandong Key Laboratory of Reproductive Medicine, Shandong Provincial Clinical Research Center for Reproductive Health, Shandong Technology Innovation Center for Reproductive Health, Jinan, Shandong, China. [3]Key Laboratory of Reproductive Endocrinology of Ministry of Education, National Research Center for Assisted Reproductive Technology and Reproductive Genetics, Shandong University, Jinan, Shandong, China. [4]CAS Key Laboratory of Genome Sciences and Information, Collaborative Innovation Center of Genetics and Development, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, China. [5]China National Center for Bioinformation, Beijing, China. [6]Nvwa life technology, Guangzhou, Guangdong, China. [7]Shanghai Key Laboratory for Assisted Reproduction and Reproductive Genetics, Shanghai, China. [8]These authors contributed equally: Yuan Gao, Lizhi Yi, Jianhong Zhan, Lijuan Wang, Xuelong Yao. ✉email: gaoyuan@sduivf.com; liuj@big.ac.cn; chenzijiang@hotmail.com

DATA AVAILABILITY

The data that support the findings in this study are deposited in the Genome Sequence Archive (GSA) for human under the accession number HRA002940. Our data are available with a signed data use agreement; please contact liuj@big.ac.cn. The downstream reuse of the controlled access datasets is restricted to non-profitable usage.

REFERENCES

1. Gardner, D. K., Lane, M., Stevens, J., Schlenker, T. & Schoolcraft, W. B. *Fertil. Steril.* **73**, 1155–1158 (2000).
2. Munné, S. et al. *Fertil. Steril.* **112**, 1071–1079.e7 (2019).
3. Jiang, L. et al. *Cell* **153**, 773–784 (2013).
4. Wang, L. et al. *Cell* **157**, 979–991 (2014).
5. Li, G. et al. *J. Genet. Genomics* **44**, 475–481 (2017).
6. Glick, I., Kadish, E. & Rottenstreich, M. *Int. J. Womens Health* **13**, 751–759 (2021).
7. Yan, J. et al. *N. Engl. J. Med.* **385**, 2047–2058 (2021).
8. Fauque, P. et al. *Clin. Epigenetics* **12**, 191 (2020).
9. Vermeiden, J. P. & Bernardus, R. E. *Fertil. Steril.* **99**, 642–651 (2013).
10. Cortessis, V. K. et al. *J. Assist. Reprod. Genet.* **35**, 943–952 (2018).
11. Court, F. et al. *Genome Res.* **24**, 554–569 (2014).
12. Zink, F. et al. *Nat. Genet.* **50**, 1542–1552 (2018).
13. Joshi, R. S. et al. *Am. J. Hum. Genet.* **99**, 555–566 (2016).
14. Jiao, X. et al. *Clin. Epigenetics* **11**, 120 (2019).

ACKNOWLEDGEMENTS

This work was supported by the grants from the National Key R&D Program of China (2022YFC2703200, 2021YFC2700600, 2021YFC2700500); Research Unit of Gametogenesis and Health of ART-Offspring, the Chinese Academy of Medical Sciences; Special Foundation of President of the Chinese Academy of Sciences; the National Natural Science Foundation of China (31871454 to J.L., 32121001to J.L.)

AUTHOR CONTRIBUTIONS

J.L., Z.C., and Y.G. conceived and designed the study. L.Y., J.Z., L.W., X.Y., J.Y., S.J., L.G., M.F., M.G., Y.Z., X.G., and K.W. conducted the experiments. L.Y. and J.Z. performed the bioinformatics analyses. J.L. and J.Z. wrote the manuscript.

COMPETING INTERESTS

X.Y. and J.L. are shareholders of Nvwa life technology. The other authors declare no competing financial interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41422-023-00809-z.

Correspondence and requests for materials should be addressed to Yuan Gao, Jiang Liu or Zi-Jiang Chen.

Reprints and permission information is available at http://www.nature.com/reprints

Supplementary information

Method

Human subjects.

The study was approved by the ethics committees at the Center for Reproductive Medicine, Shandong Provincial Hospital–Shandong University, and Beijing Institute of Genomics, reference number [2018] IRB No. (26). All the couples provided written informed consent before participation.

Clinical data collection

Study population.

We aim to enroll participants undergoing IVF cycle with at least one of PGS indications age from 20 to 45 years old. PGS indications include advanced maternal age, recurrent pregnancy loss with unknown reason, recurrent implantation failure with unknown reason, severe oligoteratozoospermia. At least two available blastocysts with morphological grade higher than BC is required for enrolled participant. Women with listed situation were not enrolled: uterine abnormalities, hydrosalpinx or any other diseases that might influence pregnancy such as type I and II diabetes, hypertension, etc.

Interventions and Duration.

Ovarian hyperstimulation (10 to 30 days): All subjects will receive controlled ovarian hyperstimulation with gonadotropin releasing hormone (GnRH) agonist long or short protocol or GnRH antagonist protocol as per physician preference.

Oocyte retrieval, ICSI and embryo culture (5-6 Days): Oocyte retrieval will be performed 34 to 36 hours after oocyte maturation trigger by hCG, GnRHa, or dual trigger. Intracytoplasmic sperm injection (ICSI) will be applied for all included subjects. On Day 5-6 of embryo culture, the morphology score of blastocyst was observed and recorded.

Embryo biopsy, verification and NGS test: Subjects in both groups will have all their embryos frozen and undergone single frozen blastocyst transfer. All blastocysts will be biopsied on Day five to six. All biopsied cells will be diagnosed by PIMS. Genome-wide DNA methylation level and chromosome ploidy will be evaluated. Euploid embryos will be transferred to uterus.

A recovery period after ovarian stimulation of 6-10 weeks to allow, shedding of the endometrial lining, decline of steroid hormones, and restoration of normal ovarian function: Endometrial preparation will be started at the second or third menses cycle after oocyte retrieval. Both natural ovulation cycle and artificial regimen will be needed for endometrial preparation.

Pregnancy evaluation and follow-up (2 weeks to 10 months): All pregnancies will be followed up till termination or delivery.

Outcome of subsequent transfers (12-21 months): If live birth is not achieved by initial transfer and there are euploid embryos, subsequent single embryo transfer(s) will be performed (up to 3 or up to one year after enrollment).

Evaluations.

Conception: After twelve to fifteen days the embryo transferred, serum Quantitative hCG will be tested. Conception will be diagnosed with a quantitative hCG of 25 mIU/ml or above. Clinical pregnancy: Twenty days after conception, transvaginal ultrasonography will be performed. Clinical pregnancy will be diagnosed with detection of an intrauterine gestational sac.

Obstetric complications, live birth, and neonatal complications including congenital anomalies: These outcomes will be determined with reference to the obstetric and neonatal medical record.

Outcomes.

The primary outcome is the live birth rate in women using the embryos with euploid chromosomes through PIMS. Secondary outcomes are the rates of pregnancy, pregnancy loss and live birth with different pattern of DNA methylation.

DNA methylation library generation and data analysis

DNA methylome preparation: DNA methylome was prepared using the optimized WGBS protocol known as 'one-tube' from our lab.[1] In brief, genome DNA of 3 to 5 cells obtained from trophectoderm of blastocyst was extracted and fragmented by sonication. The fragmented DNA was then subjected to end repair, dA tailing and adaptor ligation before bisulfite treatment. Two rounds of PCR amplification using indexed sequencing primer were then performed with 8 cycles in the first round and 8 to 12 cycles in the second round.

Sequencing read quality control and alignment: All bisulfite sequencing reads were first trimmed to remove adaptors and low quality bases using Trimmomatic (version 0.33) (parameters: LEADING:5 TRAILING:5 SLIDINGWINDOW:4:20 MINLEN:36).[2] Next, pair-end reads passed quality control were mapped to the human reference genome (hg19) using bismark (version 0.16.3) on pair-end mode (parameters: -X 1000 -non_bs_mm).[3] Single-end reads passed quality control were mapped to reference using single-end mode (parameters: -non_bs_mm). Samtools and Picard were then employed to sort and remove PCR duplication reads.[4]

DNA methylation level estimation: All sequencing reads with fully unconverted non-CpG cytosine were removed before calling DNA methylation status. To extract methylation status on cytosines, bismark_methylation_extractor was applied on pair-end reads (parameters: -p -non_overlap -comprehensive -report) and single-end reads (parameters: -s -comprehensive -report). For each CpG site, the methylation level is determined by the ratio of methylated C in covered reads to total number of reads covered this CpG. Whole genome DNA methylation level is determined by the average DNA metylation level of all covered CpG site.

Chromosomal copy number variation estimation.

The chromosome number was calculated by R package HMMcopy[5]. In brief, genome was divided into non-overlapping windows of 1Mb. Then read numbers were counted in these windows and next were corrected by GC content and mappability. The copy numbers were then segmented and classified with a robust Hidden Markov Model. The detected aneuploid embryos were excluded from embryo selection for transfer, as designed and described in the

study protocol.

Difference of methylation level between embryo with different maternal age.

Embryo with maternal age higher or less than 38 years old were separated into different groups. Embryo methylation level for each group were summarized and difference significance between groups were calculated using two-sided t test in R.

Clinical outcome statistical analysis.

To study clinical outcome differences between the embryos with different DNA methylation level, we assigned embryos into 5 different group according to their DNA methylation level. Outcomes were summarized for each group individually. Two-sided fisher exact test function in R were employed to calculate significance on difference in pregnancy rate, pregnancy loss rate or live birth rate between embryo inside and outside the methylation level window 0.25~0.27. For advanced maternal age group (maternal age ⩾38 years old) and young group (maternal age <38 years old), clinical outcome difference significance was called using same method described above. The number of windows for AMA group is decreased to 3 because the number of embryos transferred in this group is limited.

To validate the difference in outcome is not related with biased maternal age, the maternal age of embryos in the identified optimal window was compared with embryos in other windows. Two-sided T test was applied to calculate the significance of this difference.

Imprinting control regions defect analysis.

The genomic location of known ICRs were acquired from previous report[6-8]. To identify defect in ICRs, we firstly defined expected frequencies on differential methylated allele in known ICRs by extract bisulfite sequencing reads, which covered more than 3 CpGs, from all birth embryos. Allele were separated into three different methylation states, allele with DNA methylation level in 0.66~1 was denoted as 'Meth', 0.33~0.66 as 'Mixed' and 0~0.33 as 'Unmeth'. Proportion of each state in an ICR was defined as expected allelic frequencies of this ICR.

Then, we validated ICR location by our data. We merged the methylomes of birth groups, and checked the methylation levels of the reported ICRs. If the methylation levels of the reported ICRs are within 0.35 – 0.65, we regard them as the true germ line ICRs. "The true germ line ICRs" are used for further analyses in this study.

To test if a germ line ICR in a given sample is in abnormal methylation pattern, we extracted bisulfite sequencing reads of this sample in this ICR, and get the observed different methylated allele frequencies for this region using same method described above. Next, the observed allele frequencies were compared with expected allele frequencies based on allele absolute counts using fisher exact test for 'Meth' or 'Unmeth' allele. The fisher exact test P value were used to describe the ICR methylation pattern abnormalities significance with alternative hypothesis set as 'two-sided'.

Calling the differential methylated regions associated with clinical outcome.

To identify DMRs, we compared the methylomes for all transferred embryos. Human genome was split into non-overlapping sliding windows with the length of 1000 bp. The DNA

methylation level of each window in each embryo is calculated. If CpGs within a window was covered by more than 30 times in an embryo, this window would be considered as an effectively covered window for the embryo. Next, we split all embryos into two groups. Embryos producing live birth were assigned into birth group, and embryos failed to produce live birth were assigned into failure in birth group (failure-group). For each window, we calculated the proportion of the embryo number with "effectively covered window" versus total embryo number within birth group or failure-group. Low-coverage windows, defined as the proportion less than 5% in any group, were excluded from the following analyses. Finally, we calculated the mean methylation level of every window within each group, and then performed two-sided student t test between two groups. A window with different methylation level greater than 0.1 ($P<0.05$) was annotated as a DMR.

Gene annotation and statistics.

To find genes and genomic elements neighboring DMRs, NCBI RefSeq and RepeatMasker on hg19 were downloaded from UCSC table browser. Gene promoter region was defined as 1000 bp upstream and downstream of TSS. Gene has promoter overlapped with DMR with at least 1 bp was defined as TSS-DMR overlapped gene. Gene located within upstream or downstream 10 kb of DMRs were defined as DMR related gene. To analyze DMR neighboring retrotransposons, location of LINE, SINE, LTR and SVA were extracted from RepeatMasker. Retrotransposons located within upstream or downstream 1 kb of DMRs were defined as DMR related retrotransposons.

1 Li, G. *et al.* Genome wide abnormal DNA methylome of human blastocyst in assisted reproductive technology. *J Genet Genomics* **44**, 475-481, doi:10.1016/j.jgg.2017.09.001 (2017).

2 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).

3 Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571-1572, doi:10.1093/bioinformatics/btr167 (2011).

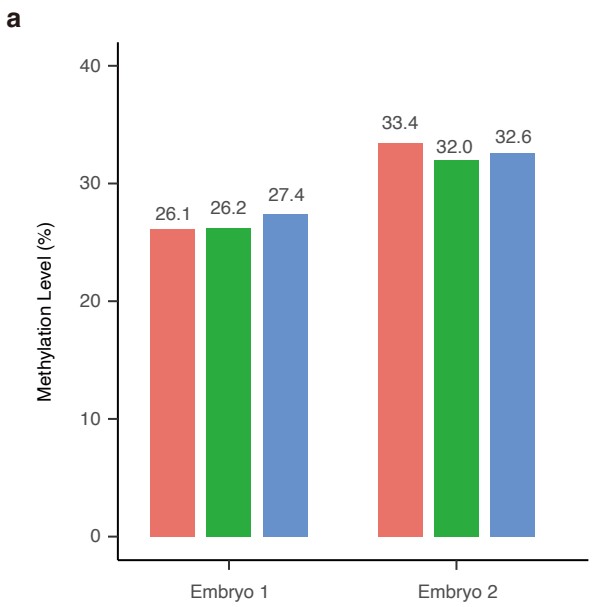
4 Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, doi:10.1093/gigascience/giab008 (2021).

5 Daniel, L., Gavin, H. & Sohrab, S. Copy number prediction with correction for GC and mappability bias for HTS data. doi:10.18129/B9.bioc.HMMcopy (2020).

6 Court, F. *et al.* Genome-wide parent-of-origin DNA methylation analysis reveals the intricacies of human imprinting and suggests a germline methylation-independent mechanism of establishment. *Genome Res* **24**, 554-569, doi:10.1101/gr.164913.113 (2014).

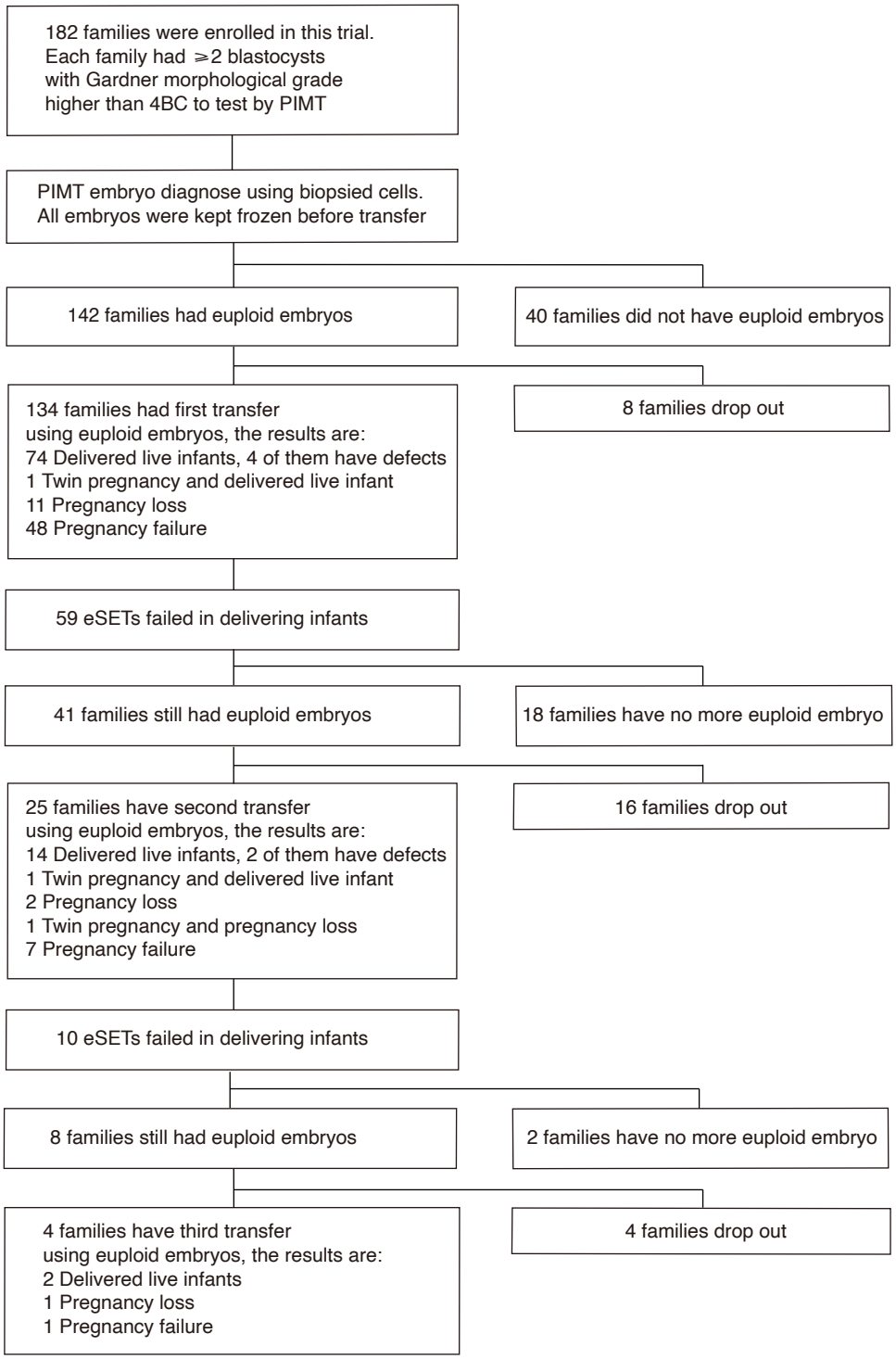
7 Zink, F. *et al.* Insights into imprinting from parent-of-origin phased methylomes and transcriptomes. *Nat Genet* **50**, 1542-1552, doi:10.1038/s41588-018-0232-7 (2018).

8 Joshi, R. S. *et al.* DNA Methylation Profiling of Uniparental Disomy Subjects Provides a Map of Parental Epigenetic Bias in the Human Genome. *Am J Hum Genet* **99**, 555-566, doi:10.1016/j.ajhg.2016.06.032 (2016).

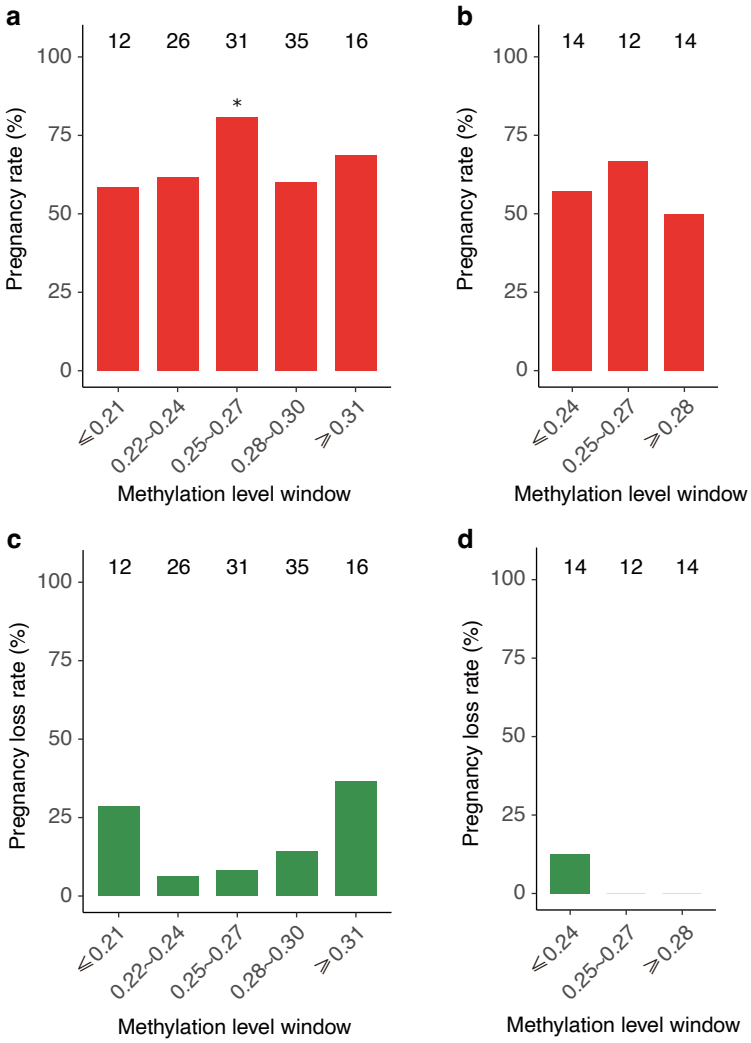


Supplementary information, Fig. S1 Methylation levels of different biopsied TE cells from the same embryo. a Three biopsies were analyzed for two embryos, respectively. DNA methylation level of each biopsy was indicated.

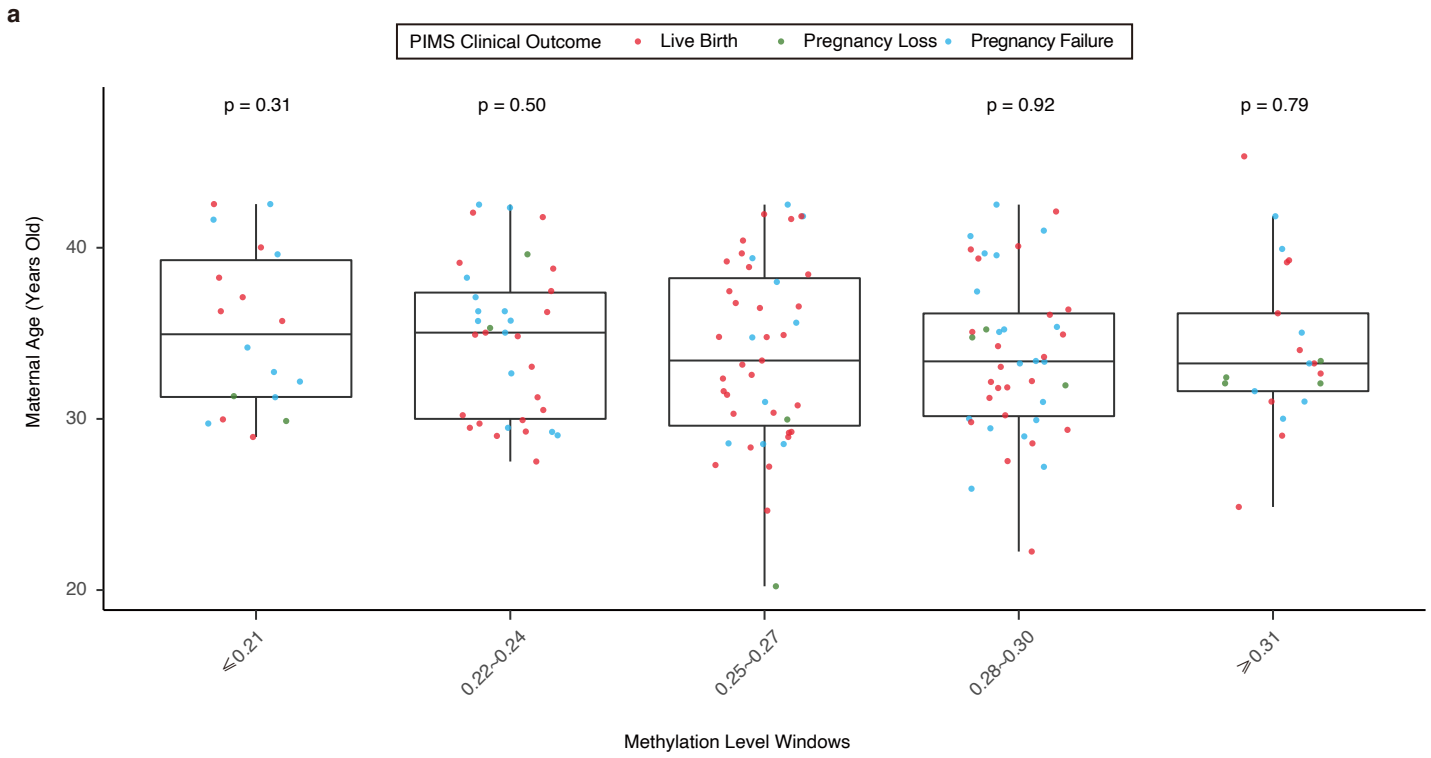
a



Supplementary information, Fig. S2 Flowchart of PIMS clinic trial. a Initially, 800 blastocysts from 182 participants were diagnosed by PIMS. Euploid frozen embryo was then selected and transferred without consideration of DNA methylation level. 48 families did not perform embryo transfer, include 40 had no euploid embryos and 8 lost to follow-up. 134 families had at least one elective single embryo transfer (eSET). In total, 163 euploid embryos performed eSETs. 3 eSETs resulted in twin pregnancy. In the following analysis, we focused in the 160 euploid embryos which produced single pregnancy or non-pregnancy. Embryo transfer were performed up to 3 times for participant when euploid embryo is available and they adhered to protocol. 3 cases of transfers resulted in twin pregnancy were excluded in downstream analysis, because assigning the DNA methylation level for these embryos are impossible.

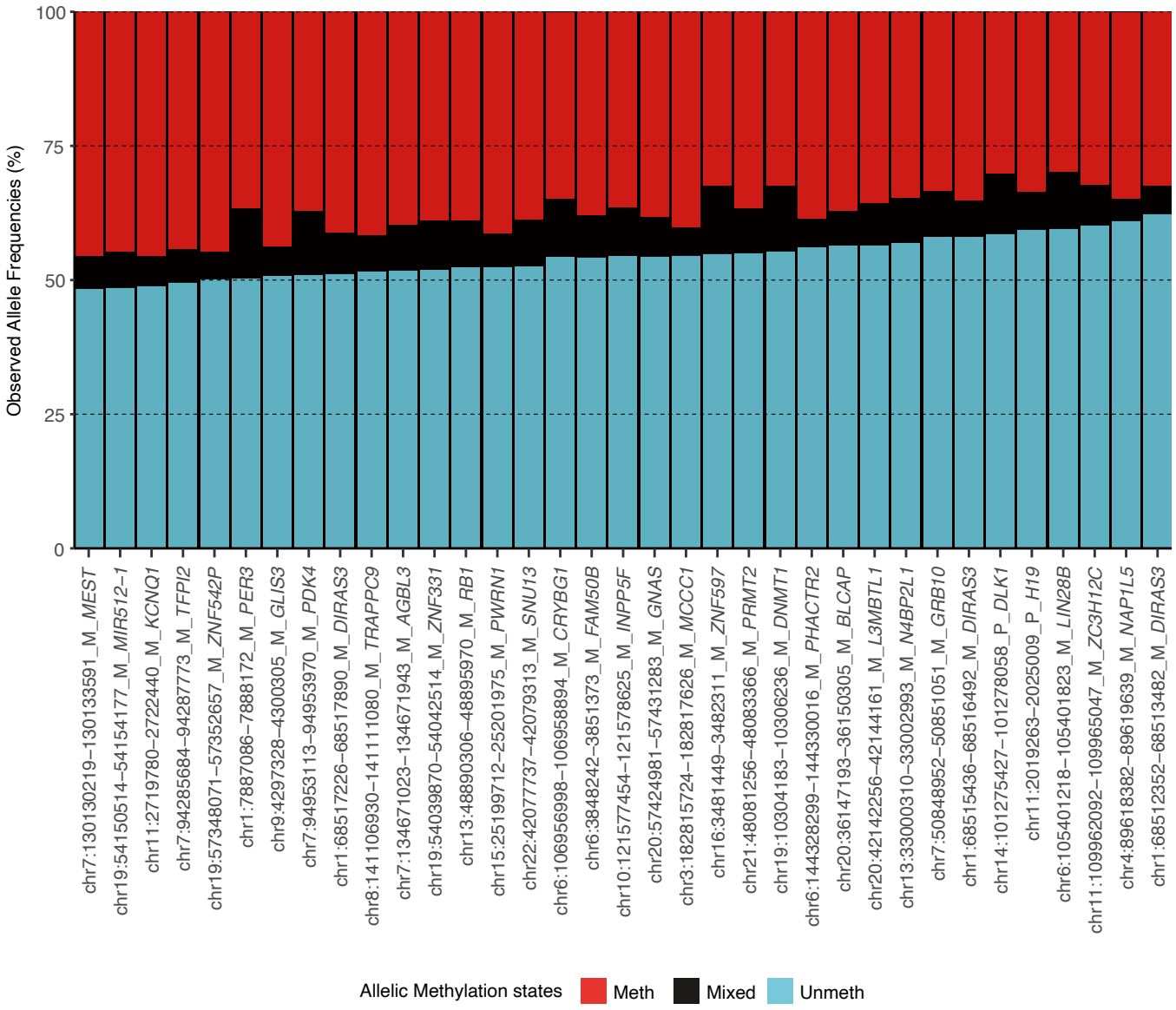


Supplementary information, Fig. S3 Pregnancy rate and pregnancy loss rate of younger and AMA women a, b Pregnancy rate of embryos with different methylation level for younger (**a**) (P=0.03) and AMA (**b**) women. **c, d** Pregnancy loss rate of embryos with different methylation level for younger (**c**) and AMA (**d**) women. * refers to $P < 0.05$ and total number of embryos transferred in each methylation level window were indicated above each columns.



Supplementary information, Fig. S4 Maternal age of embryos in different methylation window. a Each box and whisker represent the the 25th/75th percentiles and 1.5X interquartile range of maternal age of embryo in the corresponding methylation level window. Dots represent transferred embryos in PIMS. We compared maternal age between the embryos in the window of 0.25~0.27 and those in other windows, and performed two-sided T test.

a



Supplementary information, Fig. S5 Observed frequencies of differential methylated reads in germline ICRs. a Each column represents a known ICR, the proportion of different colors in each column represents the frequencies of fully-methylated (red), middle-methylated (black) and un-methylated (cyan) reads in data merged from all live birth embryos. For each ICR, their genomic location, imprinted allele (M, maternal; P, paternal) and related gene were indicated below corresponding column.